7 Foundations of Data Science (DJW)

Let $X_1, \ldots, X_{100}$ be independent samples drawn from the $\text{Exp}(\lambda)$ distribution, for some unknown parameter $\lambda > 0$.

[*Note:* The $\text{Exp}(\lambda)$ distribution has density function $f(x) = \lambda e^{-\lambda x}$, for $x > 0$. It has mean $1/\lambda$, and variance $1/\lambda^2$.]

(*a*) Show that the maximum likelihood estimator for λ is $\hat{\lambda} = 100/\sum_{i=1}^{100} X_i$.
[3 marks]

(*b*) Using the central limit theorem, find a and b such that

$$\mathbb{P}\big(1/\hat{\lambda} \in [a, b]\big) \approx 0.95$$

explaining your calculations carefully. Hence find real numbers α and β such that

$$\mathbb{P}\big(\lambda \in [\alpha\hat{\lambda}, \beta\hat{\lambda}]\big) \approx 0.95\,.$$

[6 marks]

(*c*) Explain how to use the bootstrap resampling method to approximate the probability

$$\mathbb{P}\Big(\lambda \in \big[\hat{\lambda}(1-\varepsilon), \hat{\lambda}(1+\varepsilon)\big]\Big)$$

where ε is given. In your answer, include an explanation of what is meant by 'resampling'.
[6 marks]

(*d*) Using your answer to Part (*c*), give pseudocode to compute ε such that

$$\mathbb{P}\Big(\lambda \in \big[\hat{\lambda}(1-\varepsilon), \hat{\lambda}(1+\varepsilon)\big]\Big) \approx 0.95\,.$$

Comment your code appropriately. [5 marks]